CHAPTER 21

COST CURVES

In the last chapter we described the cost-minimizing behavior of a firm. Here we continue that investigation through the use of an important geometric construction, the **cost curve**. Cost curves can be used to depict graphically the cost function of a firm and are important in studying the determination of optimal output choices.

21.1 Average Costs

Consider the cost function described in the last chapter. This is the function $c(w_1, w_2, y)$ that gives the minimum cost of producing output level y when factor prices are (w_1, w_2). In the rest of this chapter we will take the factor prices to be fixed so that we can write cost as a function of y alone, $c(y)$.

Some of the costs of the firm are independent of the level of output of the firm. As we've seen in Chapter 20, these are the fixed costs. Fixed costs are the costs that must be paid regardless of what level of output the firm produces. For example, the firm might have mortgage payments that are required no matter what its level of output.

Other costs change when output changes: these are the variable costs. The total costs of the firm can always be written as the sum of the variable costs, $c_v(y)$, and the fixed costs, F:

$$c(y) = c_v(y) + F.$$

The **average cost function** measures the cost per unit of output. The **average variable cost function** measures the variable costs per unit of output, and the **average fixed cost function** measures the fixed costs per unit output. By the above equation:

$$AC(y) = \frac{c(y)}{y} = \frac{c_v(y)}{y} + \frac{F}{y} = AVC(y) + AFC(y)$$

where $AVC(y)$ stands for average variable costs and $AFC(y)$ stands for average fixed costs. What do these functions look like? The easiest one is certainly the average fixed cost function: when $y = 0$ it is infinite, and as y increases the average fixed cost decreases toward zero. This is depicted in Figure 21.1A.

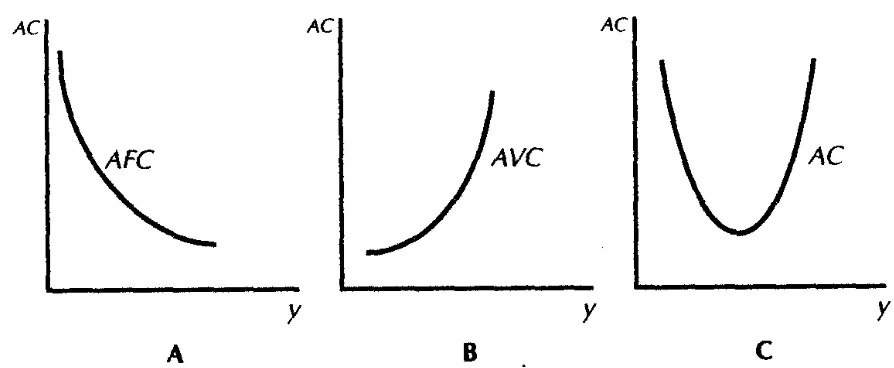


Figure 21.1 **Construction of the average cost curve.** (A) The average fixed costs decrease as output is increased. (B) The average variable costs eventually increase as output is increased. (C) The combination of these two effects produces a U-shaped average cost curve.

Consider the variable cost function. Start at a zero level of output and consider producing one unit. Then the average variable costs at $y = 1$ is just the variable cost of producing this one unit. Now increase the level of production to 2 units. We would expect that, at worst, variable costs would double, so that average variable costs would remain constant. If we can organize production in a more efficient way as the scale of output is increased, the average variable costs might even decrease initially. But eventually we would expect the average variable costs to rise. Why? If fixed factors are present, they will eventually constrain the production process.

For example, suppose that the fixed costs are due to the rent or mortgage payments on a building of fixed size. Then as production increases, average variable costs—the per-unit production costs—may remain constant for a while. But as the capacity of the building is reached, these costs will rise sharply, producing an average variable cost curve of the form depicted in Figure 21.1B.

The average cost curve is the sum of these two curves; thus it will have the U-shape indicated in Figure 21.1C. The initial decline in average costs is due to the decline in average fixed costs; the eventual increase in average costs is due to the increase in average variable costs. The combination of these two effects yields the U-shape depicted in the diagram.

21.2 Marginal Costs

There is one more cost curve of interest: the **marginal cost curve**. The marginal cost curve measures the *change* in costs for a given change in output. That is, at any given level of output y, we can ask how costs will change if we change output by some amount Δy:

$$MC(y) = \frac{\Delta c(y)}{\Delta y} = \frac{c(y + \Delta y) - c(y)}{\Delta y}.$$

We could just as well write the definition of marginal costs in terms of the variable cost function:

$$MC(y) = \frac{\Delta c_v(y)}{\Delta y} = \frac{c_v(y + \Delta y) - c_v(y)}{\Delta y}.$$

This is equivalent to the first definition, since $c(y) = c_v(y) + F$ and the fixed costs, F, don't change as y changes.

Often we think of Δy as being one unit of output, so that marginal cost indicates the change in our costs if we consider producing one more discrete unit of output. If we are thinking of the production of a discrete good, then marginal cost of producing y units of output is just $c(y) - c(y - 1)$. This is often a convenient way to think about marginal cost, but is sometimes misleading. Remember, marginal cost measures a *rate of change*: the change in costs divided by a change in output. If the change in output is a single unit, then marginal cost looks like a simple change in costs, but it is really a rate of change as we increase the output by one unit.

How can we put this marginal cost curve on the diagram presented above? First we note the following. The variable costs are zero when zero units of output are produced, by definition. Thus for the first unit of output produced

$$MC(1) = \frac{c_v(1) + F - c_v(0) - F}{1} = \frac{c_v(1)}{1} = AVC(1).$$

Thus the marginal cost for the first small unit of amount equals the average variable cost for a single unit of output.

Now suppose that we are producing in a range of output where *average* variable costs are decreasing. Then it must be that the *marginal* costs are less than the average variable costs in this range. For the way that you push an average down is to add in numbers that are less than the average.

Think about a sequence of numbers representing average costs at different levels of output. If the average is decreasing, it must be that the cost of each additional unit produced is less than average up to that point. To make the average go down, you have to be adding additional units that are less than the average.

Similarly, if we are in a region where average variable costs are rising, then it must be the case that the marginal costs are greater than the average variable costs—it is the higher marginal costs that are pushing the average up.

Thus we know that the marginal cost curve must lie below the average variable cost curve to the left of its minimum point and above it to the right. This implies that the marginal cost curve must intersect the average variable cost curve at its minimum point.

Exactly the same kind of argument applies for the average cost curve. If average costs are falling, then marginal costs must be less than the average costs and if average costs are rising the marginal costs must be larger than the average costs. These observations allow us to draw in the marginal cost curve as in Figure 21.2.

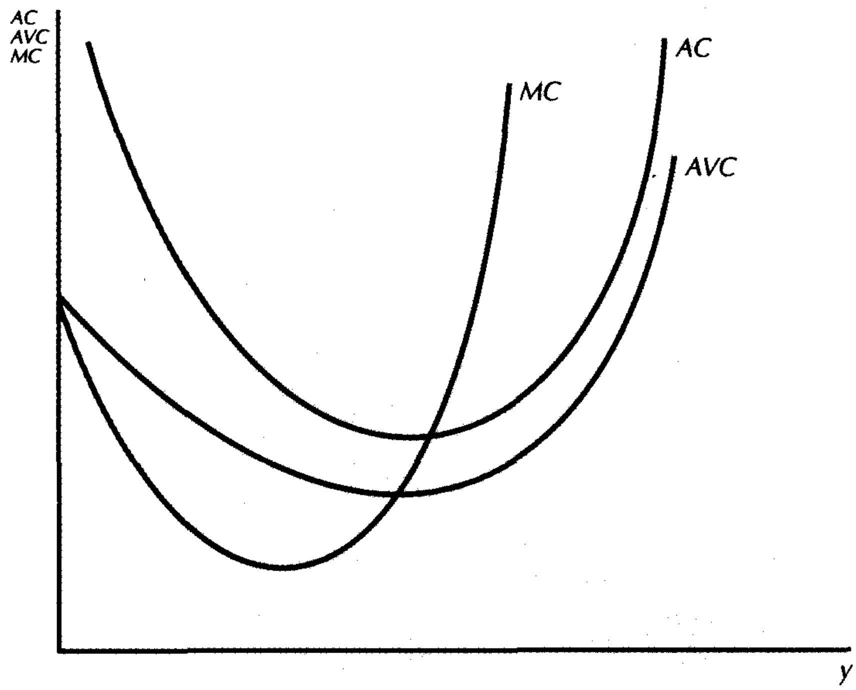
To review the important points:

- The average variable cost curve may initially slope down but need not. However, it will eventually rise, as long as there are fixed factors that constrain production.

- The average cost curve will initially fall due to declining fixed costs but then rise due to the increasing average variable costs.

- The marginal cost and average variable cost are the same at the first unit of output.

- The marginal cost curve passes through the minimum point of both the average variable cost and the average cost curves.



Cost curves. The average cost curve (AC), the average variable cost curve (AVC), and the marginal cost curve (MC).

21.3 Marginal Costs and Variable Costs

There are also some other relationships between the various curves. Here is one that is not so obvious: it turns out that the area beneath the marginal cost curve up to y gives us the variable cost of producing y units of output. Why is that?

The marginal cost curve measures the cost of producing each additional unit of output. If we add up the cost of producing each unit of output we will get the total costs of production—except for fixed costs.

This argument can be made rigorous in the case where the output good is produced in discrete amounts. First, we note that

$$c_v(y) = [c_v(y) - c_v(y-1)] + [c_v(y-1) - c_v(y-2]+ \\ \cdots + [c_v(1) - c_v(0)].$$

This is true since $c_v(0) = 0$ and all the middle terms cancel out; that is, the second term cancels the third term, the fourth term cancels the fifth term, and so on. But each term in this sum is the marginal cost at a different level of output:

$$c_v(y) = MC(y-1) + MC(y-2) + \cdots + MC(0).$$

Thus each term in the sum represents the area of a rectangle with height $MC(y)$ and base of 1. Summing up all these rectangles gives us the area under the marginal cost curve as depicted in Figure 21.3.

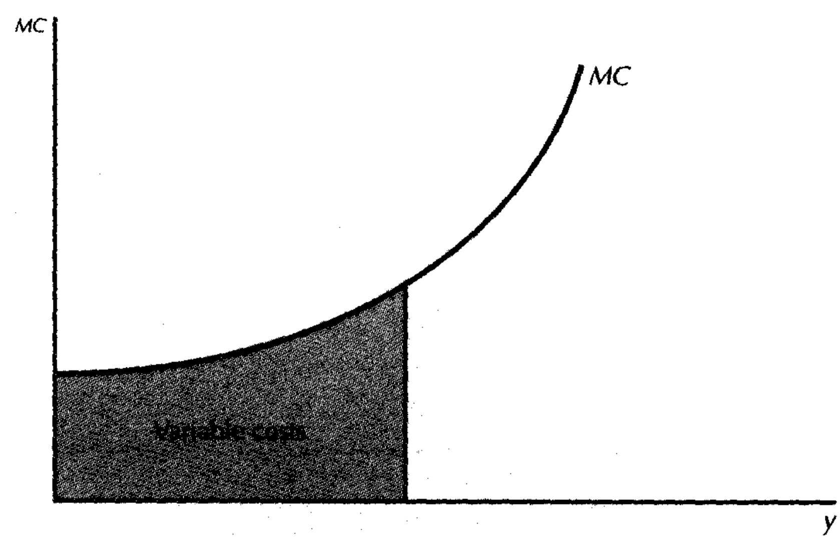


Figure 21.3 **Marginal cost and variable costs.** The area under the marginal cost curve gives the variable costs.

EXAMPLE: Specific Cost Curves

Let's consider the cost function $c(y) = y^2 + 1$. We have the following derived cost curves:

- variable costs: $c_v(y) = y^2$
- fixed costs: $c_f(y) = 1$
- average variable costs: $AVC(y) = y^2/y = y$
- average fixed costs: $AFC(y) = 1/y$
- average costs: $AC(y) = \dfrac{y^2+1}{y} = y + \dfrac{1}{y}$
- marginal costs: $MC(y) = 2y$

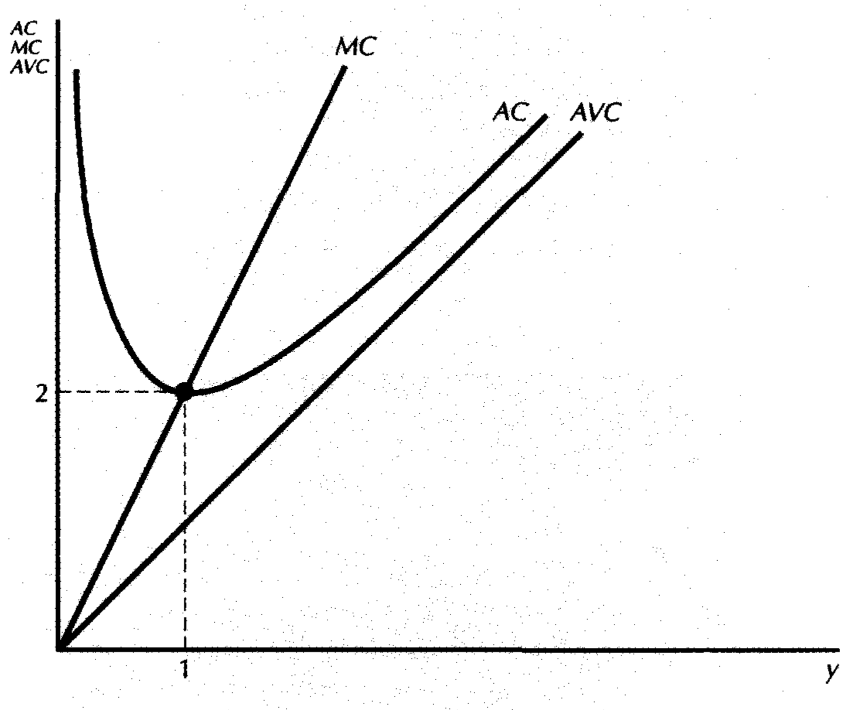
These are all obvious except for the last one, which is also obvious if you know calculus. If the cost function is $c(y) = y^2 + F$, then the marginal cost function is given by $MC(y) = 2y$. If you don't know this fact already, memorize it, because you'll use it in the exercises.

What do these cost curves look like? The easiest way to draw them is first to draw the average variable cost curve, which is a straight line with slope 1. Then it is also simple to draw the marginal cost curve, which is a straight line with slope 2.

The average cost curve reaches its minimum where average cost equals marginal cost, which says

$$y + \frac{1}{y} = 2y,$$

which can be solved to give $y_{\text{min}} = 1$. The average cost at $y = 1$ is 2, which is also the marginal cost. The final picture is given in Figure 21.4.



Cost curves. The cost curves for $c(y) = y^2 + 1$.

EXAMPLE: Marginal Cost Curves for Two Plants

Suppose that you have two plants that have two different cost functions, $c_1(y_1)$ and $c_2(y_2)$. You want to produce y units of output in the cheapest way. In general, you will want to produce some amount of output in each plant. The question is, how much should you produce in each plant?

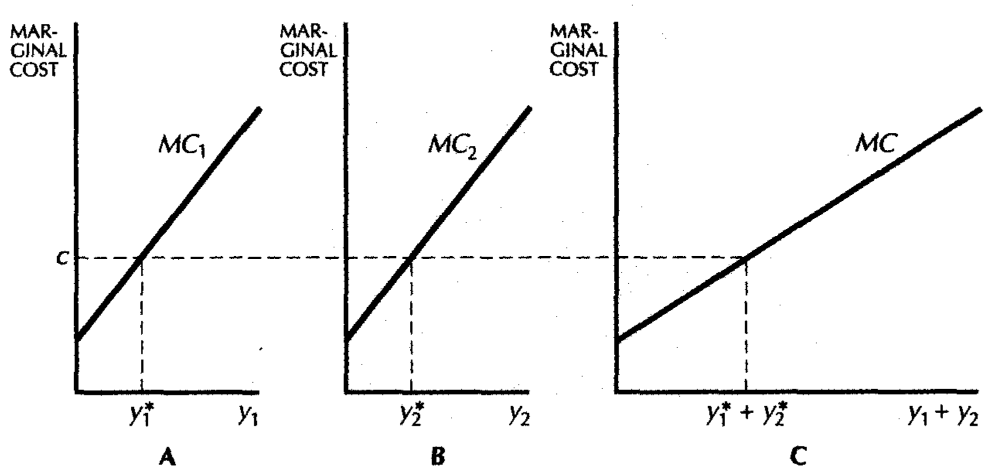
Set up the minimization problem:

$$\min_{y_1, y_2} c_1(y_1) + c_2(y_2)$$
$$\text{such that } y_1 + y_2 = y.$$

Now how do you solve it? It turns out that at the optimal division of output between the two plants we must have the marginal cost of producing output at plant 1 equal to the marginal cost of producing output at plant 2. In order to prove this, suppose the marginal costs were not equal; then it would pay to shift a small amount of output from the plant with higher marginal costs to the plant with lower marginal costs. If the output division is optimal, then switching output from one plant to the other can't lower costs.

Let $c(y)$ be the cost function that gives the cheapest way to produce y units of output—that is, the cost of producing y units of output given that you have divided output in the best way between the two plants. The marginal cost of producing an extra unit of output must be the same no matter which plant you produce it in.

We depict the two marginal cost curves, $MC_1(y_1)$ and $MC_2(y_2)$, in Figure 21.5. The marginal cost curve for the two plants taken together is just the horizontal sum of the two marginal cost curves, as depicted in Figure 21.5C.



Marginal costs for a firm with two plants. The overall marginal cost curve on the right is the horizontal sum of the marginal cost curves for the two plants shown on the left.

For any fixed level of marginal costs, say c, we will produce y_1^* and y_2^* such that $MC_1(y_1^*) = MC(y_2^*) = c$, and we will thus have $y_1^* + y_2^*$ units of output produced. Thus the amount of output produced at any marginal cost c is just the sum of the outputs where the marginal cost of plant 1 equals c and the marginal cost of plant 2 equals c: the horizontal sum of the marginal cost curves.

21.4 Long-Run Costs

In the above analysis, we have regarded the firm's fixed costs as being the costs that involve payments to factors that it is unable to adjust in the short run. In the long run a firm can choose the level of its "fixed" factors—they are no longer fixed.

Of course, there may still be quasi-fixed factors in the long run. That is, it may be a feature of the technology that some costs have to be paid to produce any positive level of output. But in the long run there are no fixed costs, in the sense that it is always possible to produce zero units of output at zero costs—that is, it is always possible to go out of business. If quasi-fixed factors are present in the long run, then the average cost curve will tend to have a U-shape, just as in the short run. But in the long run it will always be possible to produce zero units of output at a zero cost, by definition of the long run.

Of course, what constitutes the long run depends on the problem we are analyzing. If we are considering the fixed factor to be the size of the plant, then the long run will be how long it would take the firm to change the size of its plant. If we are considering the fixed factor to be the contractual obligations to pay salaries, then the long run would be how long it would take the firm to change the size of its work force.

Just to be specific, let's think of the fixed factor as being plant size and denote it by k. The firm's short-run cost function, given that it has a plant of k square feet, will be denoted by $c_s(y, k)$, where the s subscript stands for "short run." (Here k is playing the role of $\overline{x}_2$ in Chapter 20.)

For any given level of output, there will be some plant size that is the optimal size to produce that level of output. Let us denote this plant size by $k(y)$. This is the firm's conditional factor demand for plant size as a function of output. (Of course, it also depends on the prices of plant size and other factors of production, but we have suppressed these arguments.) Then, as we've seen in Chapter 20, the long-run cost function of the firm will be given by $c_s(y, k(y))$. This is the total cost of producing an output level y, given that the firm is allowed to adjust its plant size optimally. The long-run cost function of the firm is just the short-run cost function evaluated at the optimal choice of the fixed factors:

$$c(y) = c_s(y, k(y)).$$

Let us see how this looks graphically. Pick some level of output y^*, and let $k^* = k(y^*)$ be the optimal plant size for that level of output. The short-run cost function for a plant of size k^* will be given by $c_s(y, k^*)$, and the long-run cost function will be given by $c(y) = c_s(y, k(y))$, just as above.

Now, note the important fact that the short-run cost to produce output y must always be at least as large as the long-run cost to produce y. Why? In the short run the firm has a fixed plant size, while in the long run the firm is free to adjust its plant size. Since one of its long-run choices is always to choose the plant size k^*, its optimal choice to produce y units of output must have costs at least as small as $c(y, k^*)$. This means that the firm must be able to do at least as well by adjusting plant size as by having it fixed. Thus

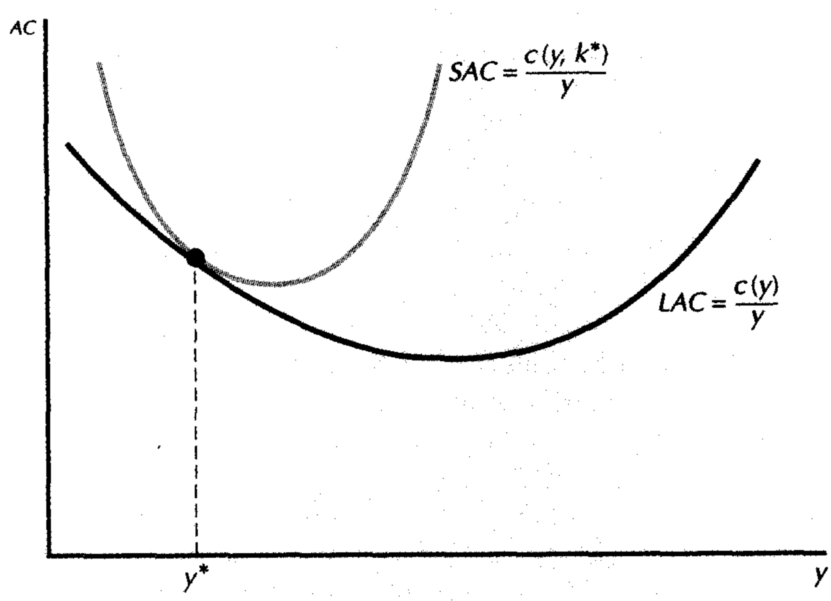
$$c(y) \leq c_s(y, k^*)$$

for all levels of y.

In fact, at one particular level of y, namely y^*, we know that

$$c(y^*) = c_s(y^*, k^*).$$

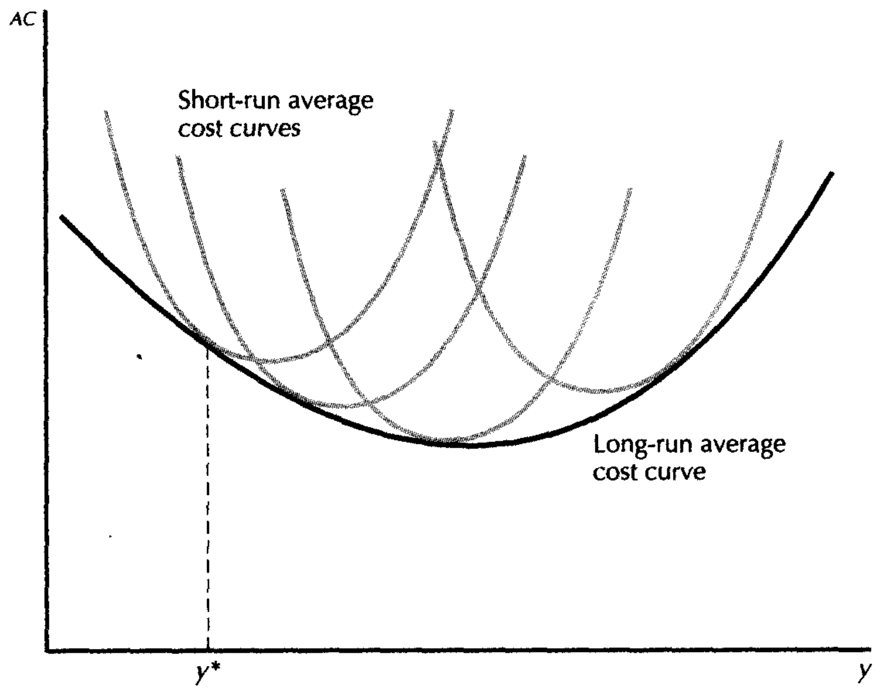
Why? Because at y^* the *optimal* choice of plant size is k^*. So at y^*, the long-run costs and the short-run costs are the same.



Short-run and long-run average costs. The short-run average cost curve must be tangent to the long-run average cost curve.

If the short-run cost is always greater than the long-run cost and they are equal at one level of output, then this means that the short-run and the long-run average costs have the same property: $AC(y) \leq AC_s(y, k^*)$ and $AC(y^*) = AC_s(y^*, k^*)$. This implies that the short-run average cost curve always lies above the long-run average cost curve and that they touch at one point, y^*. Thus the long-run average cost curve (LAC) and the short-run average cost curve (SAC) must be tangent at that point, as depicted in Figure 21.6.

We can do the same sort of construction for levels of output other than y^*. Suppose we pick outputs $y_1, y_2, \ldots, y_n$ and accompanying plant sizes $k_1 = k(y_1), k_2 = k(y_2), \ldots, k_n = k(y_n)$. Then we get a picture like that in Figure 21.7. We summarize Figure 21.7 by saying that the long-run average cost curve is the **lower envelope** of the short-run average cost curves.

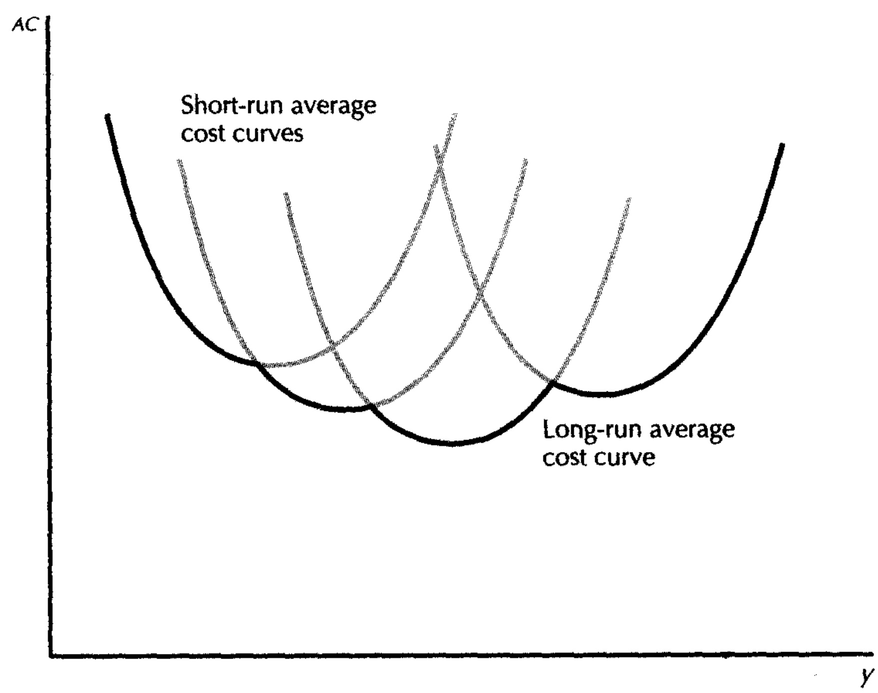


Short-run and long-run average costs. The long-run average cost curve is the envelope of the short-run average cost curves.

21.5 Discrete Levels of Plant Size

In the above discussion we have implicitly assumed that we can choose a continuous number of different plant sizes. Thus each different level of output has a unique optimal plant size associated with it. But we can also consider what happens if there are only a few different levels of plant size to choose from.

Suppose, for example, that we have four different choices, k_1, k_2, k_3, and k_4. We have depicted the four different average cost curves associated with these plant sizes in Figure 21.8.



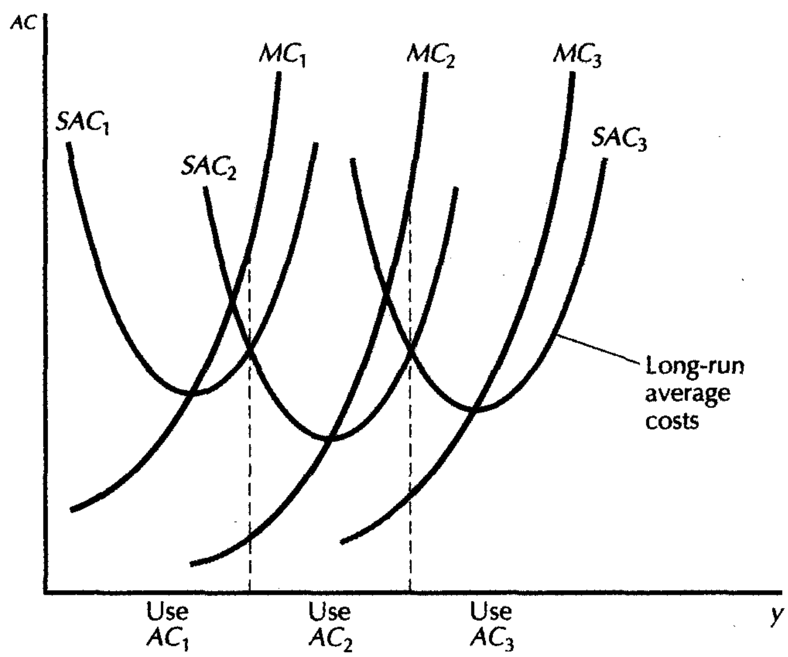
Discrete levels of plant size. The long-run cost curve is the lower envelope of the short-run curves, just as before.

How can we construct the long-run average cost curve? Well, remember the long-run average cost curve is the cost curve you get by adjusting k optimally. In this case that isn't hard to do: since there are only four different plant sizes, we just see which one has the lowest costs associated with it and pick that plant size. That is, for any level of output y, we just choose the plant size that gives us the minimum cost of producing that output level.

Thus the long-run average cost curve will be the lower envelope of the short-run average costs, as depicted in Figure 21.8. Note that this figure has qualitatively the same implications as Figure 21.7: the short-run average costs always are at least as large as the long-run average costs, and they are the same at the level of output where the long-run demand for the fixed factor equals the amount of the fixed factor that you have.

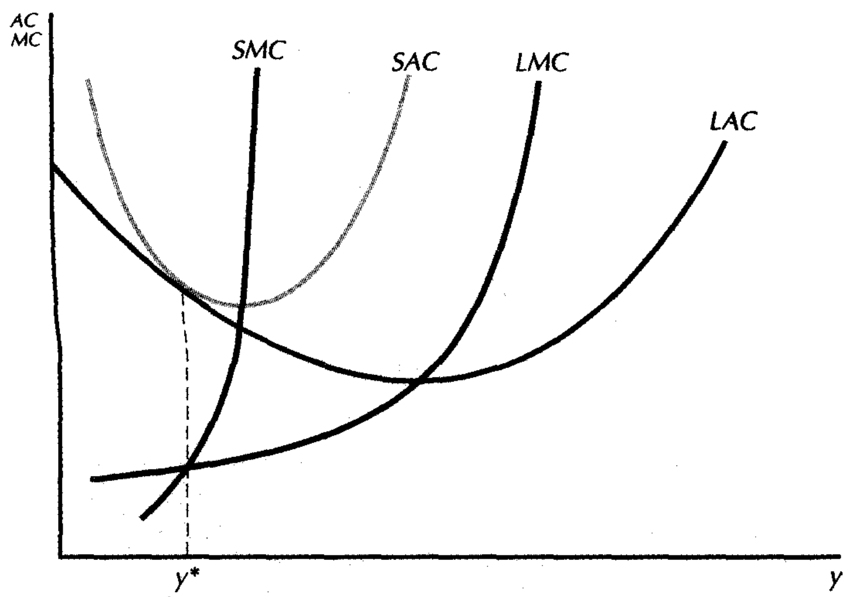
21.6 Long-Run Marginal Costs

We've seen in the last section that the long-run average cost curve is the lower envelope of the short-run average cost curves. What are the implications of this for marginal costs? Let's first consider the case where there are discrete levels of plant size. In this situation the long-run marginal cost curve consists of the appropriate pieces of the short-run marginal cost curves, as depicted in Figure 21.9. For each level of output, we see which short-run average cost curve we are operating on and then look at the marginal cost associated with that curve.



Long-run marginal costs. When there are discrete levels of the fixed factor, the firm will choose the amount of the fixed factor to minimize average costs. Thus the long-run marginal cost curve will consist of the various segments of the short-run marginal cost curves associated with each different level of the fixed factor.

This has to hold true no matter how many different plant sizes there are, so the picture for the continuous case looks like Figure 21.10. The long-run marginal cost at any output level y has to equal the short-run marginal cost associated with the optimal level of plant size to produce y.



Long-run marginal costs. The relationship between the long-run and the short-run marginal costs with continuous levels of the fixed factor.

Summary

1. Average costs are composed of average variable costs plus average fixed costs. Average fixed costs always decline with output, while average variable costs tend to increase. The net result is a U-shaped average cost curve.

2. The marginal cost curve lies below the average cost curve when average costs are decreasing, and above when they are increasing. Thus marginal costs must equal average costs at the point of minimum average costs.

3. The area under the marginal cost curve measures the variable costs.

4. The long-run average cost curve is the lower envelope of the short-run average cost curves.

REVIEW QUESTIONS

1. Which of the following are true? (1) Average fixed costs never increase with output; (2) average total costs are always greater than or equal to average variable costs; (3) average cost can never rise while marginal costs are declining.

2. A firm produces identical outputs at two different plants. If the marginal cost at the first plant exceeds the marginal cost at the second plant, how can the firm reduce costs and maintain the same level of output?

3. True or false? In the long run a firm always operates at the minimum level of average costs for the optimally sized plant to produce a given amount of output.

APPENDIX

In the text we claimed that average variable cost equals marginal cost for the first unit of output. In calculus terms this becomes

$$\lim_{y \to 0} \frac{c_v(y)}{y} = \lim_{y \to 0} c'(y).$$

The left-hand side of this expression is not defined at $y = 0$. But its limit is defined, and we can compute it using l'Hôpital's rule, which states that the limit of a fraction whose numerator and denominator both approach zero is given by the limit of the derivatives of the numerator and the denominator. Applying this rule, we have

$$\lim_{y \to 0} \frac{c_v(y)}{y} = \frac{\lim_{y \to 0} dc_v(y)/dy}{\lim_{y \to 0} dy/dy} = \frac{c'(0)}{1},$$

which establishes the claim.

We also claimed that the area under the marginal cost curve gave us variable cost. This is easy to show using the fundamental theorem of calculus. Since

$$MC(y) = \frac{dc_v(y)}{dy},$$

we know that the area under the marginal cost curve is

$$c_v(y) = \int_0^y \frac{dc_v(x)}{dx}\, dx = c_v(y) - c_v(0) = c_v(y).$$

The discussion of long-run and short-run marginal cost curves is all pretty clear geometrically, but what does it mean economically? It turns out that the calculus argument gives the nicest intuition. The argument is simple. The marginal cost of production is just the change in cost that arises from changing output. In the short run we have to keep plant size (or whatever) fixed, while in the long run we are free to adjust it. So the long-run marginal cost will consist of two pieces: how costs change holding plant size fixed plus how costs change when plant size adjusts. But if the plant size is chosen optimally, this last term has to be zero! Thus the long-run and the short-run marginal costs have to be the same.

The mathematical proof involves the chain rule. Using the definition from the text:

$$c(y) \equiv c_s(y, k(y)).$$

Differentiating with respect to y gives

$$\frac{dc(y)}{dy} = \frac{\partial c_s(y,k)}{\partial y} + \frac{\partial c_s(y,k)}{\partial k}\frac{\partial k(y)}{\partial y}.$$

If we evaluate this at a specific level of output y^* and its associated optimal plant size $k^* = k(y^*)$, we know that

$$\frac{\partial c_s(y^*, k^*)}{\partial k} = 0$$

because that is the necessary first-order condition for k^* to be the cost-minimizing plant size at y^*. Thus the second term in the expression cancels out and all that we have left is the short-run marginal cost:

$$\frac{dc(y^*)}{dy} = \frac{\partial c_s(y^*, k^*)}{\partial y}.$$